

THE GIRDEN-MURPHY PAPERS ON PK

By J. G. PRATT

INTRODUCTION

THE SEPTEMBER, 1962, issue of the *Psychological Bulletin* contained a paper entitled, "A Review of Psychokinesis (PK)," by Dr. Edward Girden (1). This article was represented by the author as a comprehensive survey of the research on PK, but it was in fact concerned only with the question of evidence for the occurrence of psychokinesis, and to the latter he gave a negative answer. In the next issue of the same periodical Dr. Gardner Murphy pointed out some of the deficiencies of Dr. Girden's review (3), and Dr. Girden added a postscript (2) in which he seemed to modify his original position in significant respects.

Ordinarily, three such papers might be accepted as disposing of the issues raised by the exchange of views. In this instance, however, there is a need to extend the discussion, particularly to make reference to the controversy in the parapsychological literature and to place it in proper perspective for students of the field. Dr. Murphy in his reply was limited by space, and he was speaking primarily as a psychologist to fellow-psychologists. He was also speaking as one who was not closely associated with the earlier research at Duke upon which the case for the occurrence of PK originally rested. He therefore directed his attention especially to the later experimental work. In doing this he pointed out Dr. Girden's shortcomings with the happy combination of gentleness and firmness that the wise teacher might use in correcting a student who has misbehaved in the expectation of being secretly applauded by his classmates.

In spite of Dr. Murphy's reply, therefore, two requirements remain to be met which justify further attention to these papers. One is the need to identify Dr. Girden's review more clearly for what it really was: an effort to discredit parapsychology by the use of illogical and unscientific tactics. The other is the need to show how Dr. Girden sidestepped the evidence with which he should have been principally concerned in dealing with the question of the strength of the case for PK. In meeting the former requirement it will be

necessary to describe the manner in which Dr. Girden approached his task as well as to deal with relevant aspects of his review. In regard to the latter, we will be chiefly concerned with the earlier evidence from "declines" in the PK data as represented by the "quarter distribution" or QD effects, the facet of the results that Dr. Murphy did not discuss fully. For the sake of completeness, the present paper will comment in turn upon all three of the papers making up this PK controversy in the *Psychological Bulletin*.

THE GIRDEN REVIEW OF PK

In his review Dr. Girden acknowledged financial support from two sources: from Stanford University through the Thomas Welton Stanford Fellowship; and from a John Simon Guggenheim fellowship. He did not indicate whether these grants ran simultaneously or whether the Stanford fellowship was awarded to extend his project. The original plan, as he stated during his one brief visit to the Parapsychology Laboratory at Duke, was to spend two years on the Guggenheim grant.

The review of PK, the first harvest from his labors, was published several years after he began his work. Dr. Girden says that this paper ". . . constitutes one phase of a larger study of the process of controversy in scientific endeavors." Can we not reasonably expect, in the light of this statement, that it should be an exemplary illustration of scientific criticism? If so, how does the performance measure up to the promise?

That Dr. Girden's approach to his task was unusual (to say the least) became apparent soon after he started his survey project. During his one-day visit to the Duke Laboratory at the beginning of his study, he so obviously did *not* ask any questions about the research that we inquired whether there was anything he wanted to learn from us. His reply was that he did not work by first gathering the facts. He wanted simply to see and talk with the people who had been active in the research in order to get a personal impression about them. He would return later for the purpose of going into the basic scientific questions about the experiments.

The second visit was never made.

As Dr. Girden continued with his project, he became increasingly evasive toward the parapsychologists regarding his plans and activi-

ties. Once during this period Dr. J. B. Rhine gave a talk at Brooklyn College while Dr. Girden was there. Girden did not attend. When Dr. Rhine was invited back another time, he agreed to go if Dr. Girden would share the rostrum to speak on his interest in parapsychology. Those who issued the invitation tried to get Dr. Girden to do so but he declined to take part in the proposed scientific debate.

Dr. Girden traveled widely, however, and he interviewed many people—especially those whom he knew to be skeptical about PK. Some of their opinions, as expressed to him either in conversation or in private correspondence, were used in his paper. But the review contains nothing that he learned at first hand from investigators who have reported significant PK results on the basis of their own research.

We are not told just when, in the course of his study of parapsychology, Dr. Girden decided to begin his parapsychological publications with a "review" of PK. But he did so decide, and the reasons he gives are interesting—especially in relation to his obvious (but nowhere clearly stated) purpose and his way of going about his task. He says:

There are a number of reasons to justify a review devoted specifically to PK. After study of the published data and discussions with interested parties here and abroad, it seems clear that all of the issues which have been raised with respect to ESP also appear in connection with PK. The topic constitutes a unit which can be considered within the limits of a single publication. It is an area with which the academic psychologist is generally unfamiliar. Although criticism of some of the PK reports, especially the earlier dice tests, have appeared, there has appeared no assessment of psychokinesis as a whole.

This quotation is as near as Dr. Girden comes to stating what he intended to do; the reader is left to find out for himself that the review is *concerned with the literature on psychokinesis solely from the point of view of evidence for the occurrence of PK.*

This would be appropriate, of course, as the scope of a review, but the taking of this limited objective without clearly stating such an intention or acting in a manner consistent with it is entirely inappropriate. An objective limited to the question of the occurrence of PK obviously means that Dr. Girden should concern himself

solely with the *best* evidence to see whether any of it stands up to critical scrutiny, not with *all* the experimental reports, and especially not with the *weakest* PK evidence he can find. Yet he writes in the above paragraph as if he needs to survey everything on PK: "The topic constitutes a unit which can be considered within the limits of a single publication . . . assessment of psychokinesis as a whole." This idea that his self-made assignment was to cover *all* the PK literature, when he was really only concerned about whether the evidence justifies the conclusion that PK occurs, led him into writing an article that is entirely inappropriate and misleading. Indeed, the inconsistency between *purpose* and *performance* is so glaring that one cannot help wondering if Dr. Girden himself was unaware of it.

The simple declarative statement that PK ". . . is an area with which the psychologist is generally unfamiliar" is a strange and, if true, revealing announcement. Why should the author of a general review on a topic that is of fundamental concern for psychology assume that psychologists are uninformed about it? Even if he discovered in the course of his study that there is this unusual degree of ignorance about PK research, why should he simply note the fact in passing, as if this were a normal and proper state of affairs? The PK literature is as available to psychologists generally as it was to Dr. Girden. Is he saying that psychologists have been guilty of neglecting an area that obviously should be of great concern to them? And if so, is he implying that his review is going to provide the information they need to help them decide about PK? Or is he only paying lip service to his obligations as a critic? Has he prejudged the issue unfavorably and does he think that the psychologists likewise have such strongly adverse views about the evidence for PK that they will agree with his negative opinions regardless of whether he presents them logically and informatively? If the latter is the case, his review is more of an indictment of himself and of the psychological profession than it is of parapsychology!

The Girden review, both by its style and its contents, clearly shows that the author started his task already committed to an unfavorable verdict against PK. Nowhere does he show the slightest interest in any question beyond that of whether there is evidence that PK occurs. Yet he uses the method of casting his net over the

whole literature of the field and then looking critically at his haul to see if he has caught anything unfit for scientific consumption. This approach is obviously not appropriate for dealing with the question of whether the *best* evidence available establishes PK. His question should be: Is there *any* evidence which adequately supports the positive conclusions of the PK research workers? To answer this question, he does not need to examine everything that he can sweep up in his net. Especially, he may not select out the worst experimental specimens he can find (without regard to the reservations that the investigators themselves expressed) and claim that their shortcomings weaken or destroy the better evidence for PK.

What Dr. Girden obviously should have done was to examine the evidence upon which those who have reached a positive conclusion regarding the reality of PK have based their case. If he could actually show that this best evidence is fallacious, his case against PK would be won. This he clearly does not do, and the impression is inescapable that he knew he was avoiding the issue and attempting, instead, to win a hopeless case by cleverly introducing irrelevant side issues.

The best evidence for PK is found in the internal decline effects, especially the data contained in the three studies of the "quarter distribution" (QD) effect (4, 5, 6). Dr. Girden did not miss this important point. To avoid the necessity of frankly facing this evidence, he carefully attempted to discredit it in the minds of his "uninformed" psychological readers. Early in the article he speaks, in passing, about the QD studies as a "post-mortem evaluation." Having thus devaluated it, he puts off facing this central issue as long as possible. Near the end of his article, in a section on declines, when he can no longer avoid the issue, Dr. Girden sweeps aside the QD evidence on the basis of assertions which he does not bother either to justify or to relate to the facts. He says: ". . . since the [decline in scoring] hypothesis was derived from these data, independent evidence is required for its confirmation." This dictum ignores the fact that the QD hypothesis was formulated on the basis of the analysis of the *first* series examined out of the 18 included in the study of the QD of the page. The other 17 series therefore properly constituted a confirmation of the hypothesis. To argue, as Dr. Girden does, that no confirmation of the hypothesis was possible

because the data were already in existence at the time is like arguing that no science of geology is possible because it is based upon "post-mortem" hypotheses formulated on the basis of data already in existence on earth!

But let us give Dr. Girden the benefit of the doubt regarding his not recognizing that the first QD paper was a confirmation of a hypothesis formulated on the basis of only one out of 18 series used in that study. There is still no reason why he should have missed the important point that the *third* paper in the series, the QD of the half-set, was a confirmation made on a predictive basis in the strictest sense of the word. To make this point quite clear, I will spell out the facts.

I was absent from the Duke Laboratory on war work when the QD effect was discovered, and I had not at any time examined any of the original PK data. Thus I could not have formed any impression about the distribution of the hits in the records in units of the data structure smaller than the page and the set, the units studied in the QD analyses that had been published during my absence. In the fall of 1943 I returned to the Laboratory for two months. During that time I carried out a re-analysis of the PK data upon which the QD studies were based, primarily to check upon the accuracy of the analyses. Before this recheck was begun, I suggested that the opportunity could be used to serve a second purpose as well: an independent confirmatory study of the QD effect. This could be achieved by tabulating the hits so as to make possible a test of the diagonal decline within the half-set. By restricting this study to those series in which the half-sets fell *entirely within* one of the four quarters of the page, the findings of the proposed new investigation would be statistically independent of those obtained in the original QD analysis.

We made clear in our report on this study that the investigation was regarded as confirmatory on a strictly predictive basis. Eight series were available for this analysis. The difference between the hits in the upper left-hand and the lower right-hand quarters of the half-sets was evaluated for each series by the CR method. When these differences were squared and combined as chi-square, they gave a total chi-square indicating that the overall level of significance was $P = .000,005$. This study is only a part of the evidence from

declines that Dr. Girden bypasses. It is singled out for special attention here because it so clearly makes ridiculous his attempt to use his description of the QD analysis as a "post-mortem" evaluation as an excuse for ignoring this evidence.

But even after he had thus summarily dismissed the early QD evidence, he could not leave the topic of declines alone. He discussed the Gibson Machine Series adversely on the ground that the experimenter had used the higher faces as targets more often on the left-hand side of the page and the lower faces more often on the right-hand side. Since the dice had excavated spots, the faces with higher numbers were lighter. This probably led to a higher rate of hitting in the upper left quarter of the page than in the lower right quarter. But this is no reason for Dr. Girden's statement generalizing the defect of this one experiment to cover, as he does, the PK data as a whole: ". . . the position effect most likely was an attribute of dice bias, i.e., associated mainly with high faces."

Rhine and Humphrey clearly indicated in the first QD paper that they were aware of the fact that some series (including the Gibson Machine Series) did not meet the requirements for an even distribution of the six faces of the die across the record page, and they showed that the QD effect of the page was still highly significant when all these series were omitted from the analysis (4, p. 42).

There are other unwarranted attempts by Dr. Girden to discredit some of the PK evidence that he could not criticize legitimately. Some of the best evidence from individual experimental series he attempts to discount by disparagingly referring to the subjects as "sensitives." Other work (which was not reported as conclusive evidence of PK) he derides as "solo" (that is, unwitnessed) investigations. In fact, one would never know from reading his review that the research workers all along made a careful distinction between exploratory research (which could not by itself produce crucial evidence for PK regardless of how significant the results turned out to be) and confirmatory experiments in which the conditions were adequate to eliminate all of the counterhypotheses to PK. For him, all the experiments not measuring up to the highest standards of crucial proof of PK are subject to being cited as examples of critical flaws in the research. He attacks such work regard-

less of the fact that the authors of the original reports did not draw final conclusions from their results.

Dr. Girden's treatment of his material simply does not communicate any coherent information to the uninformed reader. His sentences fall into four main categories: 1. Those which are fragments of factual information so incomplete or misleading as to convey little meaning to the reader. (For example: "The total die throws were 1,440 with Fisk and 1,392 with West, but the authors . . . concluded that a significant difference between them 'has hardly been established.'") 2. Those which are flat assertions, given without logic or justification, regarding what one would do and would not do in making an experimental test of the PK hypothesis. ("Of a number of interesting considerations, it is self-evident that the most elementary requirements necessitated the equal representation of all six faces as targets in some randomized order and the tabulation of all die faces on all trials." "The prior analysis in terms of target scores readily suggests that the disclosed weaknesses in experimental design apply with equal force to the lawful decline in scoring.") 3. Those giving personal evaluations without any effort at justification, as if the author hopes, by expressing his own prejudices against PK, to strike a responsive chord in his readers. ("None of these crucial weaknesses in experimental design is rectified by the post mortem report of a significant decline in hits in these already completed studies.") 4. And finally, remarks or statements quoted out of context from the literature under review and generalized in a wholly unwarranted manner to apply to the PK evidence as a whole. (A striking example of this sort is to be found on p. 378 of Dr. Girden's article, half-way down the second column, where he quotes lines from an article of mine in which I was pointing out that the targets of the Gibson Machine Series were not randomly distributed across the page. Dr. Girden makes it appear that this quotation applied to all the early dice tests and that it thus invalidated the decline evidence as a whole!)

The total effect is comparable to a court trial in which Dr. Girden has personally assumed the roles of the indicting grand jury, the prosecuting attorney, the judge, and the trial jury. Could the confidence he seemed to feel have been, if really genuine, due to the fact that he staged his indictment and trial of PK and handed down

his pre-determined adverse verdict two months before the counsel for the defense, Dr. Gardner Murphy, was to present his case?

THE MURPHY REPORT ON GIRDEN'S REVIEW

Dr. Murphy showed by a few well chosen examples how Dr. Girden's treatment of the PK literature was in most instances grossly misleading. This he accomplished with telling effect, in spite of his great gentleness. His remarks were devastating even though he was addressing his corrective statements to scientific colleagues who could not be assumed to be informed about the area of research under consideration and even though they had been hopelessly confused by the tangled web of Dr. Girden's twisted arguments. The points which Dr. Murphy drives home in his masterful rebuttal are the following: 1. Dr. Girden under-rates the degree of confirmation of the earlier Duke evidence for PK obtained by research workers in other centers. 2. He wrongly interprets the experimental procedures followed, or unjustifiably claims that the weaker conditions of the preliminary or exploratory stage of an investigation invalidates the later, more crucial stage. 3. He is wrong in his statement, offered with increasing emphasis as he nears the end of his review, that the PK experiments were not concerned with a real psychological hypothesis. 4. His insistence upon the need for an experimental design that involves a statistical test of the difference between wishing and non-wishing trials is simply not permissible. 5. The PK experimenters were acutely aware of the matter of dice bias (particularly as it is found in using dice with excavated spots), and they were not, as Dr. Girden implies, tripped up by their failure to take account of this factor in the design of their experiments. 6. Dr. Girden is in error when he says that an unsuccessful repetition of an experiment that formerly gave statistically significant results cancels out the original findings and that a series of solo experiments by an investigator somehow nullifies a later two-experimenter research specifically designed to see if the earlier results can be confirmed. 7. It is unreasonable to demand, as Dr. Girden does, that controls be achieved in terms of perfection of dice and apparatus rather than in terms of procedural controls that nullify the effects of physical bias. 8. Dr. Girden's phraseology is unnecessarily hostile in some instances. 9. His dismissal of the

outstanding success on the initial trials of repeated short subseries of throws as occurring where the "physical conditions were likely to be the unsteadiest" is unsupported, whereas there is no doubt of a *psychological* difference between the first and subsequent trials.

THE GIRDEN POSTSCRIPT

Dr. Girden starts his rebuttal by calling attention to the fact that Dr. Murphy called his own paper a "report" on the Girden review. Apparently Girden recognized that this was an unusual characterization; but apparently he did not recognize that his review was a phenomenon that called for a *report*—certainly it did not need a *reply*! To attempt to "reply" to it would have been pointless, while in a "report" one could more properly characterize it for what it was.

Dr. Girden complains about the difficulties of making a detailed presentation of an unfamiliar subject for a professional audience. He partly blames the editor for the shortcomings of his review. This practically amounts to an admission of the futility of his presentation. He then attempts to escape from the predicament into which he has gotten himself by quoting Dr. Murphy as speaking out strongly regarding the importance of finding repeatable experiments in parapsychology. But what, one may ask, does the voicing of an ideal which might some day be attainable in this pioneering branch of science have to do with the evaluation of the best of the evidence for PK that is available now, during the early stage of the research?

Finally, he ends with a milder statement than his earlier negative attitude would lead one to expect: "Concerning the existence of PK, this writer has no strong opinion pro or con but, on the basis of the available evidence, the soundest judgment is a Scottish verdict: *not proven*." So he closes with an admission that the evidence is balanced: too strong to disregard; too weak to permit a final positive verdict. This is not a bad state of affairs for the PK hypothesis, but it is not a good way to leave the case as far as psychology is concerned. If the first concentrated attack upon psychokinesis has left, in the view of the critic, even a first-base case for PK, this should be welcomed as a challenge to pursue research on the problem.

REFERENCES

1. GIRDEN, EDWARD. A review of psychokinesis (PK). *Psychol. Bull.*, 1962, **59** (No. 5), 353-88.
2. ———. A postscript to "A Review of Psychokinesis (PK)." *Psychol. Bull.*, 1962, **59** (No. 6), 529-31.
3. MURPHY, GARDNER. Report on paper by Edward Girden on psychokinesis. *Psychol. Bull.*, 1962, **59** (No. 6), 520-28.
4. RHINE, J. B., and HUMPHREY, BETTY M. The PK effect: Special evidence from hit patterns. I. Quarter distributions of the page. *J. Parapsychol.*, 1944, **8**, 18-60.
5. ———. The PK effect: Special evidence from hit patterns. II. Quarter distributions of the set. *J. Parapsychol.*, 1944, **8**, 254-71.
6. RHINE, J. B., HUMPHREY, BETTY M., and PRATT, J. G. The PK effect: Special evidence from hit patterns. III. Quarter distributions of the half-set. *J. Parapsychol.*, 1945, **9**, 150-68.

Parapsychology Laboratory
Duke University
Durham, North Carolina